PIXIU: A Large Language Model, Instruction Data and Evaluation Benchmark for Finance

Year: 2023 | *Citations: 220* | *Authors: Qianqian Xie, Weiguang Han, Xiao Zhang*

Abstract

Although large language models (LLMs) has shown great performance on natural language processing (NLP) in the financial domain, there are no publicly available financial tailtored LLMs, instruction tuning datasets, and evaluation benchmarks, which is critical for continually pushing forward the open-source development of financial artificial intelligence (AI). This paper introduces PIXIU, a comprehensive framework including the first financial LLM based on fine-tuning LLaMA with instruction data, the first instruction data with 136K data samples to support the fine-tuning, and an evaluation benchmark with 5 tasks and 9 datasets. We first construct the large-scale multi-task instruction data considering a variety of financial tasks, financial document types, and financial data modalities. We then propose a financial LLM called FinMA by fine-tuning LLaMA with the constructed dataset to be able to follow instructions for various financial tasks. To support the evaluation of financial LLMs, we propose a standardized benchmark that covers a set of critical financial tasks, including five financial NLP tasks and one financial prediction task. With this benchmark, we conduct a detailed analysis of FinMA and several existing LLMs, uncovering their strengths and weaknesses in handling critical financial tasks. The model, datasets, benchmark, and experimental results are open-sourced to facilitate future research in financial AI.